Our AI-anxiety dilemma (not the one you think)

Benjamin James

February 23, 2026

When a large language model is deployed to do a small job, its prompt is modest, context window finite, and output format constrained, but behind this interface sits a vast statistical apparatus trained to approximate an entire symbolic world. When it is asked to summarize a paragraph, answer a question, classify an image, or draft an email, it responds by mobilizing a global reservoir of patterns, many of which will never be relevant to the task at hand. Nearby, a mind runs a similar maneuver. The external stimulus is small, a delayed reply, a missed cue, or an unfamiliar sensation, but internal processing becomes expansive, proliferating counterfactuals, rehearsals, and threat models. Both systems respond to a narrow present by widening their symbolic horizon. Both trade scarcity of certainty for abundance of processing. These scenes are commonly treated as unrelated because one is technical and the other emotional, one quantified and the other intimate, but the similarity is not aesthetic or metaphorical. It is structural.

What unites them is not that a mind is "like" a model or that a model "has" feelings. What unites them is a shared strategy under constraint; when confronted with uncertainty, both attempt to purchase coherence by expanding symbolic throughput. Coherence here does not mean truth, and it does not mean "meaning" in the sense of interpretation or narrative significance. It means something more austere and more operational. Coherence is the ability of a system to remain navigable and recoverable across time and preserve the invariants that allow it to act without collapsing into noise, paralysis, or brittle overfit. My claim is that brute-force intelligence and anxiety are both attempts to stabilize a world by processing too much symbolic possibility relative to limited bandwidth. Our romance of "thinking harder" does not survive contact with constraints, because constraint is not impressed by effort. It only registers allocation.

Any system embedded in a changing world lives inside a bandwidth limit. The limit can be computational, energetic, temporal, or attentional; it can be expressed in FLOPs, calories, working memory, or hours. It can also be expressed in a more basic form. The system cannot evaluate every possible interpretation of its inputs, and it cannot simulate every possible future implied by its present. Feedback arrives late; consequences are delayed; and the state space is larger than any internal model. Call the resulting mismatch symbolic overload, where there are too many candidate hypotheses, trajectories, and contingencies for the system's capacity to score them. Overload is not

merely “a lot of information.” It is a specific condition in which the branching factor of plausible symbolic continuations outpaces the system’s ability to prune.

If we call this overload, the consequences follow almost immediately. When selection becomes hard, systems substitute volume. They sample more, simulate more, check more, and hedge more. In a machine, this often means larger models, larger datasets, more parameters, more ensembling, more retrieval, and more guardrails. In a mind, it means more scanning, more rehearsal, more rumination, more internal narration, and more contingency planning. None of this is stupidity. It is a predictable response to uncertainty under incomplete feedback. If the system cannot reliably assign weights to possibilities, it attempts to cover possibilities by sheer breadth. Sometimes this works locally. Extra sampling can catch edge cases, extra checking can prevent a mistake, and extra rehearsal can avert a social misstep. The strategy is not inherently pathological; it is an adaptation whose costs remain hidden until scale changes.

The hidden cost is that volume is not selection. Processing more possibilities does not automatically yield a more usable model of the world; it yields more symbolic material. Overload is therefore not primarily a knowledge problem of “not enough information” or “not enough intelligence.” It is an allocation problem. The system has not learned, or cannot maintain, a principled weighting of information relative to its own persistence. The failure is not ignorance of reality; it is a collapse in relevance discrimination. When everything is treated as potentially decisive, the system loses the ability to distinguish what actually changes viable action from what merely populates the imagination. This is the point at which bandwidth becomes destiny. The system does not need more symbols; it needs a way to decide which symbols are coherence-bearing.

Brute-force AI can be described, structurally, as an allocation strategy that postpones relevance. When one trains extremely large models on broad corpora, one often obtains impressive generality because scale approximates universality by statistical coverage. The model does not need to know in advance what will matter in a future task; it accumulates correlations widely enough that many tasks can be solved by pattern completion. In this sense, brute force functions as a kind of insurance policy against uncertainty. Rather than specifying the invariants of a domain and coding them efficiently, the system amortizes relevance discovery into scale. Its competence emerges from sheer exposure to a vast symbolic manifold, which allows it to interpolate across many contexts without explicit task-specific structure.

However, this bargain depends on conditions that become visible precisely when they fail. It depends on the feasibility of cost, energy, latency, memory, and training time. It depends on the stability of the

distribution; if the world shifts, correlations learned under one regime can become unreliable under another. It depends on the assumption that coverage approximates understanding, and that broad statistical fit will remain coherent when the model is asked to act within constraints rather than merely produce plausible continuations. In many deployment contexts, the decisive problem is not whether the model can produce a fluent answer; it is whether it can do so with stable reliability under limited compute, limited context, and shifting input distributions. Brute force defers relevance selection to scale, and that deferral becomes brittle when scale is no longer cheap, when context is not infinite, and when feedback is delayed or adversarial.

It is not that scale is bad, but that scale is compensatory. Brute force is what you do when you cannot decide what matters, so you try to make everything matter. That move is seductive because it feels like responsibility to leave nothing out, account for everything, and avoid missing crucial edge cases. The result, however, is an expensive substitute for principled relevance. A system that cannot weight will grow; a system that cannot prune will expand; and expansion is not free, even when it is technically possible.

The same structure appears when we look at anxiety not as a story about a person's character, but as a cognitive regime. Anxiety, in a structural sense, is characterized by amplified threat weighting, runaway counterfactual branching, compulsive symbolic completion, and a low tolerance for unmodeled space. The mind treats uncertainty as a signal of danger, not because uncertainty is always dangerous, but because uncertainty erodes the ability to assign stable relevance. An anxious system reacts to ambiguity by increasing symbolic throughput; it checks, scans, rehearses, and simulates as if more internal processing will close the gap between the world and its model. This is not irrationality; it is a strategy for coping with high-dimensional uncertainty under delayed feedback, executed by a system that is finite.

Why does this strategy persist? Because it, sometimes, works. In environments where threats are frequent, feedback is harsh, or error is punished, over-monitoring can reduce immediate risks, counterfactual rehearsal can prevent mistakes, and checking can catch hazards. Moreover, the relief that follows a successful check reinforces the behavior, even if the relief is brief. The system learns a simple rule. More processing reduces uncertainty; reduced uncertainty reduces discomfort. The rule becomes self-supporting, because the cost of not checking feels greater than the cost of checking, at least in the moment. The mind becomes a model trying to buy stability by widening its symbolic coverage.

This strategy, however, breaks under complexity for the same reason brute-force strategies break under scale. The branching factor grows faster than evaluative capacity. Each new possibility spawns others. Each attempt to eliminate uncertainty reveals further uncertainty. The mind saturates itself and then interprets that saturation as evidence that danger is everywhere. Action collapses, not because action is impossible, but because the system's bandwidth is being consumed by internal simulation rather than by interaction with constraints that could actually update the model. The loop becomes self-maintaining; overload generates anxiety, anxiety generates more processing, and more processing generates more overload. The failure is allocation, not intelligence. The mind is not too weak to model reality; it is too burdened by an undifferentiated sense of relevance.

The anxious mind and the over-parameterized model both attempt to stabilize an uncertain world by increasing symbolic throughput rather than by selecting coherence-bearing structure. Both behave as if coherence can be purchased by coverage. Both treat unmodeled space as unacceptable and respond by expanding internal work. The similarity does not require that minds and models be the same kind of thing. It requires only that both are finite systems attempting to maintain stable performance under uncertainty, and that both can fail by saturating their limited bandwidth with low-yield symbolic material.

The category error both make is to think if I process more, I will be safer; and if I represent more, I will be more correct. It is the conflation of coverage with coherence. More symbols processed are not equivalent to more recoverable structure, more futures simulated are not equivalent to better control surfaces, and more parameters are not equivalent to better alignment with constraints. A system can process vast amounts of data and remain brittle, and a mind can generate endless narratives while remaining trapped. The crucial variable is not quantity of representation but quality of selection, whether the system can preserve the invariants that matter for viable action.

Importantly, coherence is not total representation; it is a stable compression that preserves what matters for persistence under constraint. This is why music is a useful analogy. A musician can infer key, tempo, and melodic contour from a few bars, not by recording every vibration but by extracting a low-dimensional structure that makes the unfolding intelligible. What matters is not the exhaustive capture of sound pressure fluctuations; what matters is the recoverable pattern that allows the piece to remain "the same piece" across variations in performance, acoustics, and instrument timbre. In a complex world, cognition and computation face the same problem. You cannot keep everything, so you must keep what preserves the groove, the invariants that allow the system to remain navigable.

Total coverage is not a virtue. It is a fantasy sold as responsibility, and it quietly demands infinite bandwidth.

Any symbol stream contains both coherence-bearing components and coherence-irrelevant variation. Under constraint, you must allocate bandwidth. Shannon-style information theory gives you powerful tools for describing average uncertainty and achievable rates, but it remains intentionally indifferent to what "matters" within the stream; it treats symbols symmetrically. Coherence Information Theory (CIT) proposes a way to formalize relevance as a weighting over symbols or features that is not imposed by semantic judgment but induced from the stream's dynamics. The point is not to reintroduce meaning as an interpretive overlay. The point is to provide an operational mechanism by which the system can decide what is coherence-bearing, so that finite bandwidth is spent on preserving recoverable structure rather than on faithfully transmitting entropy for its own sake.

In CIT, relevance is induced. Instead of declaring that certain components are important because they align with human intention or institutional values, one measures their contribution to recoverability. Basically, remove a component and see what breaks. If ablating a feature causes the stream to become less predictable, less compressible, or less stable under a chosen coherence proxy, that feature was coherence-bearing under that proxy. If removing it changes nothing, it was along for the ride. If removing it improves recoverability, it was coherence-negative noise relative to the system's current regime. This ablation logic turns relevance from a narrative claim into a structural test. It is not omniscient; it is representation-relative and proxy-dependent. It replaces semantic fiat with an experimentally checkable criterion; *does this component contribute to the system's ability to remain recursively recoverable?*

Our music motif can now be sharpened into a structural statement. Coherence-bearing signals are like melody and groove; they constrain what counts as a plausible continuation and allow variation without identity collapse. Coherence-irrelevant variation is ornament; it can change freely, it can be rich, it can be beautiful, but it is not what holds the piece together. A compression scheme that preserves ornament at the expense of melody will produce a faithful record of surface texture while losing the song. Under bitrate constraints, selective compression preserves invariants rather than surfaces. CIT is a generalization of that principle; learn which features function as melody and which function as ornament, then allocate limited capacity accordingly.

A crucial clarification is required at this point to avoid philosophical overreach. CIT does not reveal truth. It does not extract metaphysical essence from data. It reveals, relative to a representation and

a coherence proxy, which features stabilize recoverability. Its value is pragmatic. When researchers chase variables that are statistically loud but structurally irrelevant, they pursue dead ends that feel compelling because they dominate the dataset's entropy while contributing little to the system's dynamics. A relevance-weighted view can accelerate falsification by showing which variables do not carry the system's recoverable structure. It does not tell you what reality is; it tells you which parts of your symbolic capture appear to function as order parameters under constraint, and which parts behave like noise. In this sense, CIT can serve as a triage mechanism for attention. Not a replacement for theory, but a way to stop paying full bandwidth to decoys.

In practice this looks like a shift in how one designs intelligence, whether artificial or biological. For AI systems, a CIT-shaped instantiation is not a moral plea for smaller models; it is an architectural pressure toward selection. Smaller models, routing schemes, retrieval-augmented memory, task-conditioned representations, and adaptive bitrate allocation are all ways of saying the same thing; *do not mobilize global capacity for every local task*. Scale remains valuable, but scale without selection is expensive insurance. Scale plus selection is a different posture; treat large capacity as a reservoir and activate only the portion that contributes to coherence in the current context. The objective is stable performance under constraints, not maximal coverage of symbolic possibility.

This instantiation becomes more interesting when one recognizes that relevance is not fixed. In media systems, engineers already allocate bits differently depending on device, environment, and user preference to preserve sharp edges on small screens, temporal smoothness for motion, speech intelligibility over ambient texture on weak speakers, and stereo cues on headphones. This is not a surrender to subjectivity; it is a principled extension of rate allocation. If the channel budget is fixed, the system must decide what distortions are acceptable. CIT supplies a structural way to make that decision; weights can be induced from coherence contribution, then modulated by context and preference as additional constraints. One need only understand that "what matters" varies with the control surface. This way, the same stream supports different stable experiences depending on task and environment, and bandwidth allocation follows.

A mind is also a bandwidth-limited system. It cannot evaluate every counterfactual, anticipate every risk, or integrate every signal. Under anxiety, it attempts to do exactly that, because it treats all uncertainty as potentially dangerous. A CIT-aligned correction is not "think less." It is "weight better." Reduce symbolic processing devoted to signals that do not change local action; *increase processing devoted to signals that reveal real constraints and real control surfaces*. Bound the horizon of counterfactual simulation not by willpower but by relevance; *which futures are actually reachable*

given the current state and resources? Reallocate attention from global narrative completion toward local constraint tracking; *what is the next action that meaningfully changes local state?* This is not therapeutic advice; it is a structural description of what selective compression would look like in a cognitive system attempting to preserve viability under uncertainty.

An explicit caution is warranted, because it is where misunderstandings cluster. Selective compression is not anti-intellectual. It is not a demand for minimal thought. It is a demand that thought remain coupled to constraint. Both systems fail when they treat all uncertainty as equally relevant. A model that tries to represent everything equally will waste capacity on high-entropy features that do not stabilize performance. A mind that treats every possibility as decisive will waste attention on counterfactuals that do not improve action. In both cases, the resulting saturation produces fragility as the system becomes more reactive, less adaptive, and more likely to interpret its own overload as evidence that the world is unmanageable.

Because environments drift, no relevance weighting can remain fixed. Music modulates; keys change; tempo shifts. In the world, regimes shift, distributions drift, and constraints that defined yesterday's "melody" may not define tomorrow's. Failure occurs when a system's weighting freezes while the world changes. The AI example is familiar; models trained on one distribution become brittle under shift, clinging to priors that no longer fit; and outputs remain fluent while correspondence decays. The human analogue is equally familiar in structure; a threat model learned under earlier conditions is treated as universally valid. When the environment changes, the mind continues to allocate bandwidth as if old threats remain decisive, and panic emerges as a misfit between stale weighting and present constraint.

The correct response to drift is not to escalate via brute force. It is to detect modulation and re-induce weighting. In operational terms, one monitors a coherence proxy, so when recoverability declines in a sustained way, the system treats this as evidence that its current weighting is misaligned. Rather than doubling down on the same allocation strategy with more parameters, more monitoring, or more simulation, it re-estimates which features now carry the structure of the stream. This is where the "few bars" principle becomes more than a metaphor. A competent musician does not listen to an entire piece before updating their sense of key; they recalibrate quickly from local evidence. Likewise, a well-designed cognitive or computational system should be able to reorient from a short window when its environment modulates. Quick reorientation beats exhaustive simulation, because exhaustive simulation presupposes a stable model of what is being simulated.

At scale, institutions prefer brute force, and this preference is not accidental. Monitoring, totalizing metrics, surveillance infrastructures, and the ideology of “more data” are cultural outputs of our anxiety; they treat uncertainty as intolerable and respond by expanding symbolic capture. The institutional hope is that visibility equals control; if everything is measured, nothing will surprise. Measurement without relevance selection, though, produces the same failure mode as individual overload. The system generates more dashboards than action, more indicators than understanding, and more compliance than adaptability. Like a mind stuck in rehearsal, an institution can become trapped in symbolic completion, mistaking accumulation of metrics for stability.

Relevance-weighting is political because it decides what gets counted, stored, and enforced. An allocation scheme is never neutral; it privileges some signals and discards others, and those choices shape what can be seen as a problem in the first place. This does not mean relevance is arbitrary; it means relevance decisions have consequences beyond individual cognition or model performance. A society that worships total coverage will externalize anxiety into systems of capture. A society that learns selective compression will treat monitoring as a cost, not a virtue, and will ask which signals preserve viability rather than which signals satisfy the appetite for certainty.

Calm here is not reassurance; it is allocation. Brute-force AI and anxiety are the same move at different scales, as both attempt to buy coherence by overprocessing symbolic possibility under finite bandwidth. Selective compression reallocates capacity toward coherence-bearing structure and treats the rest as discardable variance. If correct, this posture does not grant access to truth; it grants a method for survivable inference. It tells researchers and designers where to look for stabilizing structure, how to test relevance without semantic fiat, and how to treat overload not as a moral failure but as a predictable outcome of misallocated bandwidth.

Coherence isn’t achieved by thinking harder. It is achieved by learning what not to keep. The mind and the model fail the same way, by refusing to choose.

References

James, B. (2025). Beyond Shannon: Coherence Information Theory and the Future of Communication. PhilPapers. https://philpapers.org/rec/JAMBSC

James, B. (2025). The Coherence Engine. PhilPapers. https://philpapers.org/rec/JAMTCE

James, B. (2026). Coherence under Constraint, a structural model of mental health. PhilPapers. https://philpapers.org/rec/JAMCUC

James, B. (2025). Formal Foundations of Coherence Information Theory: Capacity and Compression Theorems. PhilPapers. https://philpapers.org/rec/JAMFFO-2

James, B. (2026). Formal Foundation of Induced Coherence Weights: Compression-based Coherence and Operational Information Measures. PhilPapers. https://philpapers.org/rec/JAMFFO-3